How does the ventral pathway contribute to spatial attention and the planning of eye movements?

Fred H. Hamker

California Institute of Technology, Division of Biology 139-74,
Pasadena, CA 91125, USA
fred@klab.caltech.edu
http://www.klab.caltech/∼fred.html

Abstract. Cortical organization of vision appears to be divided into two pathways: the ventral pathway and the dorsal pathway. Models of vision have generally adopted this separation into a functional division such that recognition is supposed to be located in the ventral pathway and spatial attributes are processed in the dorsal pathway. I suggest a less distinct separation. According to my model the ventral pathway contributes to the selection of the location of an object by feedback connections. Those projections localize the object of interest by transferring information about its features in IT to cells with smaller receptive fields in V4 and earlier. I demonstrate the performance of the model in a visual search task which demands an eye movement towards a target.

1   Introduction

Visual perception is proposed to rely on a pathway for object vision, the "what" pathway and one for spatial vision, the "where" pathway [1]. A refinement of this concept emphasized the relevance of the "where" pathway for action control [2]. Almost all computational models of visual perception and attention follow this separation between "where" and "what". The general idea is, that the dorsal pathway first selects the location of an object and then the ventral pathway recognizes it by analyzing only a spatially defined part of the scene [3]. This decoupling of recognition and selection has the advantage of a facilitated recognition as compared to a fully parallel approach, since it is not practicable to apply several object models at the same time at several locations [4]. However, such a model of perception has its limitation if we search for a specific object. How could the "where" pathway know what is relevant?
The relevance of an object seems to be reflected by the activity of IT cells [5] [6]. Although the initial activation of IT neurons is largely stimulus driven and cells encoding target and non-target become activated, different populations compete for representation and typically the cells encoding the non-target are suppressed. Such competition is assumed to be biased by top-down feedback from working memory [5] [6]. A computational approach by Usher and Niebur [7] shows that

a parallel competition based on lateral interactions is sufficient to qualitatively replicate some of those findings, but they argue that the parallel stage is useless in case of a search for a conjunction and the decision has to be based on a serial scan of all objects.

It was suggested that the frontal eye field (FEF) could implement a saliency map by the convergence of information from different brain areas [8]. This raises the question how the FEF knows what is task relevant and where the object of interest is located. The FEF has connections to occipital, temporal and parietal areas, the thalamus, superior colliculus and prefrontal cortex [9]. The projections from V2 and V3 are weak, from V4 intermediate and heavy from TEO. Anterior IT cortex does not project directly to FEF. Information about the target features could be received from prefrontal areas and compared with features of intermediate complexity from V4 and TEO. This would require that the FEF or related areas perform a match detection in topological and topographic space. Alternatively, Desimone and Duncan [10] speculate "at some point in time, mechanisms for spatial selection may also be engaged to facilitate localization of the target for the eye movements". Some authors proposed feature specific top-down influences [11] [12] that could guide attention before the eye movement is planned. However, their implementation and exact function remained mysterious. Others suggested a top-down directed beam within the ventral pathway [13]. Only recently the influence of top-down feedback is beginning to be investigated more closely [14] [15] [16] [17] [18] [19]. In this paper I suggest that the visual areas process incoming stimuli first in a parallel bottom-up manner without a significant bottleneck and then acquire a more detailed knowledge about an object of interest by feedback. I show that such feedback within the ventral pathway can account for goal directed covert and overt search. Even for conjunction search a serial scan is not imperative.

2   Model

I model aspects of the areas V4, IT, FEF and PF and refer to the model by the prefix M (Fig. 1). M-IT, M-V4 and M-PF are subdivided into different dimensions (e.g., color and shape). My model consists of ascending populations, called (s) stimulus cells that can be primed by feedback connections and descending populations (t) target cells that project the dominant patterns back into the source areas.

The model prefrontal cortex serves for two major functions, memorizing a pattern in M-PFwm (working memory) cells and indicating a match of the incoming pattern with the memorized pattern in M-PF match cells. Thus, M-IT cells can only drive M-PFm cells when their pattern matches the prior knowledge from M-PFwm cells.

The neurons in the FEF can be categorized based on their responses to visual stimuli or to saccade execution into visual, visuomovement, fixation and movement cells [20]. I consider (v) visuomovement, (f) fixation and (m) movement cells in my model (Fig. 1). The M-FEFv neurons receive convergent afferents

from features in M-V4 at the same retinotopic location and add-up across all dimensions. M-FEFf cells generally inhibit M-FEFm cells. A threshold detection of the M-PF match cells is applied to determine if the target is in the search array. In this case the input into the M-FEFf cell is removed and thus the mapping from sensory to motor is facilitated. M-FEFv cells activate M-FEFm cells by surround inhibition. Since there is evidence that saccades are elicited when movement related activity in the FEF reaches a particular level [21], I assume a fixed threshold in M-FEFm cells to initiate a saccade. A spatially organized gain control input of M-V4 and M-IT stimulus cells originates from from M-FEFm cells.

M-PFwm cells modulate visual processing via feedback into M-ITs according to the current goal of the task. The resulting local increase of firing in M-ITs cells is directed further downwards by feedback form M-ITt cells to M-V4s cells. Thus, increased local activity in M-V4 enhances the visually responsive neurons in the frontal eye field, such that these cells reflect the task-relevance of a location.
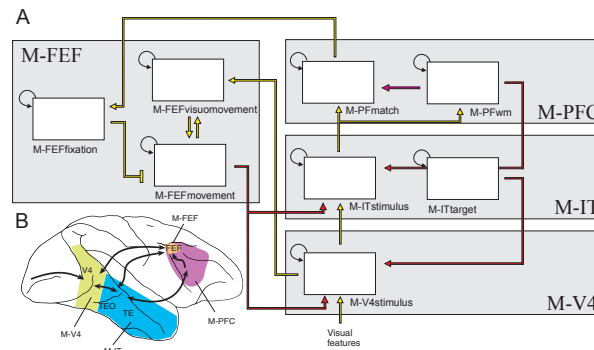

Fig. 1. (A) Sketch of the simulated areas. Each box represents a population of cells. The activation of those populations is a temporal dynamical process. Bottom-up (driving) connections are indicated by a bright arrow and top-down (modulating) connections are shown as a dark arrow. (B) Outline of the minimal set of interacting brain areas. Our model areas are restricted to elementary but typical processes and do not replicate all aspects of these areas.

3 Results

In order to demonstrate the possible role of feedback in the ventral pathway I simulated a memory guided search task [6] (Fig. 2A). If the same cued object reappears in the search array, the condition is called 'Target Present'. In the 'Target Absent' condition the cue stimulus is different from the stimuli in the choice array. Now a saccade has to be withheld.

The target was presented to the model and its features have been memorized in M-PFwm cells. Prior to the onset of the search array the active M-PFwm cells

increase the baseline activity of the M-IT cells selective for the target (Fig. 2B). When the search array appears, inputs are processed bottom-up without any strong bottleneck. Each cell initially encodes the presence of its preferred stimulus, but the target cell shows an early advantage due to top-down modulation from M-PFwm cells. Between 150 and 300 ms the cells encoding the non-target get suppressed although the input is still present, whereas the cells encoding the target remain active. A crucial condition is the target absent condition. Both non-targets decrease their activity, but less than in the distractor suppression case. A simple winner-take-all competition would not replicate the experimental data because due to noise in the system, a non-target would be selected in the target absent condition. My simulation results even match the temporal course of activity of IT cells in the different conditions of the experiment from Chelazzi et al. [6]. This constraint allows me to give reliable predictions of the processing in other areas.

The model predicts that the early advantage of IT cells encoding the target is sent to V4 cells, which have smaller RFs and creates an early target effect in V4 (see also [15]). Recent cell recordings confirmed this prediction: During the early phase until 150 ms after array onset, V4 cells show a slight target effect, which is stronger when two stimuli are located within a V4 receptive field [6].

Since FEFv neurons receive their main input from M-V4 an enhancement within the topographic/topological(feature) space is transferred into topographic space, such that a target selection is possible. This result explains how the visual cells of the FEF might discriminate over time the target from the distractor in conjunction visual search. The advantage in different dimensions adds up. The location of the target receives enhanced input from both dimensions. Locations encoding distractors sharing a single feature with the target receive enhanced input just from one dimension. The temporal course of activity of the FEFv and FEFm cells is similar to what has been found in experiments [8] [24]. FEFm cells quickly discriminate the target from the non-target.

The fronal eye field and areas within the dorsal pathway form a fronto-parietal network. These areas can use such a discrimination for overt and covert search. In overt search an eye movement is executed when the activity of the FEF movement cell reaches a threshold. Covert search is possible if activity, e.g. from the movement cells, reenters extrastriate visual cortex and enhances the input gain in V4 and IT in a spatially organized manner.

4 Discussion

This study demonstrates how findings in single cell recordings can be used to constrain models of perception. Each modeled area exhibits a temporal course of activity that has been observed by similar physiological experiments performed by various investigators. What are the major findings and predictions of this study for modeling object recognition and attention? First of all, the ventral pathway encodes an object of interest as well as its location. The model predicts that one role of feedback is to enhance the gain of cells encoding features of the
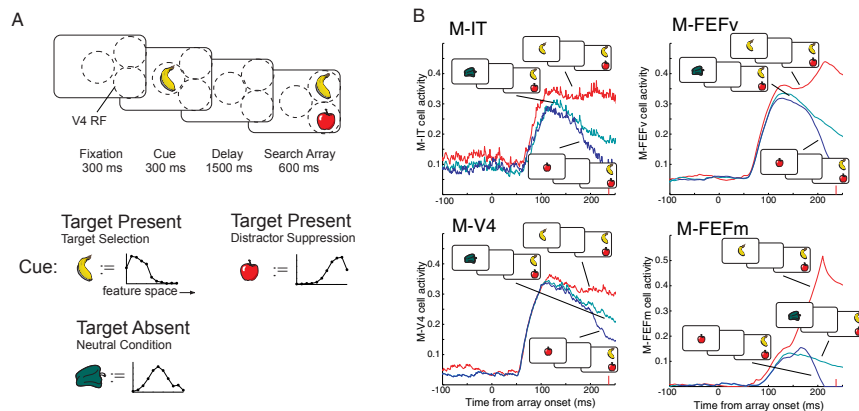
Fig. 2. (A) Simulation of the experiment of Chelazzi et al. [6]. The objects are represented by a noisy population input, here illustrated by a snapshot. RF's without an object just have noise as input. Each object is encoded within a separate RF, illustrated by the dashed circle, of M-V4 cells in two simulated dimensions (only one is shown). All M-V4 cells are within the RF of the M-IT cell population. The model has to indicate a successful search, by selecting the previously shown object as the target of an eye movement. (B) Activity within the model areas aligned to the onset of the search array in the different conditions.

object of interest. Such a mechanism would allow for a foreground-background discrimination throughout the ventral pathway down to V1.

Second, object recognition and attention recruit the same neural architecture. Recognition is related to the firing of detector cells and attention is typically implemented by control units. My model does not contain any control units. Competition and cooperation within the recognition network implements a dynamic filter that allows the brain to connect planning processes with the physical world. As a result, suppressive and facilitatory effects occur, commonly referred to as "attention".

References

1. Mishkin, M., Ungerleider, L.G., Macko, K.A.: Object vision and spatial vision: Two cortical pathways. Trends in Neurosci. **6** (1983) 414–417.
2. Milner, A.D., Goodale, M.A.: Visual pathways to perception and action. Progress in Brain Research **95** (1993) 317–337.
3. Koch, C., Ullman, S.: Shifts in selective visual attention: towards the underlying neural circuitry. Human Psychology **4** (1985) 219–227.
4. Ballard, D.H, Brown, C.M.: Principles of animate vision. In: Aloimonos, Y. (eds.): Active Perception. Lawrence Erlbaum Associates (1993) 245–282.

5. Chelazzi, L., Miller, E.K., Duncan, J., Desimone, R.: A neural basis for visual search in inferior temporal cortex. Nature **363** (1993) 345–347.
6. Chelazzi, L., Duncan, J., Miller, E.K., Desimone, R.: Responses of neurons in inferior temporal cortex during memory-guided visual search. J. Neurophysiol. **80** (1998) 2918–2940.
7. Usher, M., Niebur, E.: Modeling the temporal dynamics of IT neurons in visual search: A mechanism for top-down selective attention. J. Cog. Neurosci. **8** (1996) 311–327.
8. Thompson, K.G., Bichot, N.P., Schall, J.D.: From attention to action in frontal cortex. In: Braun, J., Koch, C., Davis, J.L. (eds.): Visual Attention and Cortical Circuits. MIT Press, Cambridge (2001), 137–157.
9. Schall, J.D., Morel, A., King, D.J., Bullier, J.: Topography of visual cortex connections with frontal eye field in macaque: Convergence and segregation of processing streams. J. Neurosci. **15** (1995) 4464–4487.
10. Desimone, R., Duncan, J.: Neural mechanisms of selective attention. Annu. Rev. Neurosci. **18** (1995) 193–222.
11. Treisman, A., Sato, S.: Conjunction search revisited. J. Exp. Psychol. Hum. Percept. Perform. **16** (1990) 459–478.
12. Wolfe, J.: Guided search 2.0 A revised model of visual search. Psychonomic Bulletin & Review **1** (1994) 202–238.
13. Tsotsos, J.K., Culhane, S.M., Wai, W., Lai, Y., Davis, N., Nuflo, F.: Modeling visual attention via selective tuning. Artificial Intelligence **78** (1995) 507–545.
14. Koechlin, E., Burnod, Y.: Dual population coding in the neocortex: A model of interaction between representation and attention in the visual cortex. J. Cog. Neurosci. **8** (1996) 353–370.
15. Hamker, F.H.: The role of feedback connections in task-driven visual search. In: Heinke, D., Humphreys, G.W., Olson, A. (eds.): Connectionist Models in Cognitive Neuroscience. Springer Verlag, London (1999), 252–261.
16. Hamker, F.H.: Distributed competition in directed attention. In: Baratoff, G., Neumann, H. (eds.): Dynamische Perzeption, Proceedings in Artificial Intelligence, Vol. 9. AKA, Akademische Verlagsgesellschaft, Berlin (2000) 39–44.
17. van der Velde, F., de Kamps, M.: From knowing what to knowing where: modeling object-based attention with feedback disinhibition of activation. J. Cogn. Neurosci. **13** (2001) 479–491.
18. Corchs, S., Deco, G.: Large-scale neural model for visual attention: integration of experimental single-cell and fMRI data. Cereb. Cortex **12** (2002) 339–48.
19. Roelfsema, P.R., Lamme, V.A., Spekreijse, H., Bosch, H.: Figure-ground segregation in a recurrent network architecture. J. Cogn. Neurosci. **14** (2002) 525–537.
20. Schall, J.D., Hanes, D.P., Thompson, K.G., King, D.J.: Saccade target selection in frontal eye field of macaque. I. Visual and premovement activation. J Neurosci **15** (1995) 6905–6918.
21. Hanes, D.P., Schall, J.D.: Neural control of voluntary movement initiation. Science **274** (1996) 427–430.
22. Hamker, F.H. Attention as a result of distributed competition. Soc. Neurosci. Abstr., Vol. 27, Program No 348.10, 2001.
23. Chelazzi, L., Miller, E.K., Duncan, J., Desimone, R.: Responses of neurons in macaque area V4 during memory-guided visual search. Cereb Cortex **11** (2001) 761–772.
24. Bichot, N.P., Rao, S.C., Schall, J.D.: Continuous processing in macaque frontal eye cortex during visual search. Neuropsychologia **39** (2001) 972-982.